Trust, but Verify: An Improved Estimating Technique Using the Integrated Master Schedule (IMS)

Eric Lofgren – Technomics, Inc.
NASA Cost & Schedule Symposium
August 29-31, 2017

> *"**Trust, but verify** is a form of advice given which recommends that while a source of information might be considered reliable, one should perform additional research to verify that such information is accurate, or trustworthy. The original Russian proverb is a short rhyme which states, Доверяй, но проверяй (doveryai, no proveryai)."*

It has long been the wonder of management why the Integrated Master Schedule (IMS) fails to give advanced warning of impending schedule delays. An estimator may follow authoritative guidance in analysis of schedule health using key metrics, supposing that such checks authenticate schedule realism. Why, then, do practitioners find themselves caught off guard by slips when their IMS appears in good health? Answers to this question follow from observing the evolution of the IMS over the course of its submissions. By independently tracing activities in an IMS, this paper will show that early performance to baseline is a good measure of realized slip and provides the framework for a new estimating technique.

Part I will motivate for the paper by with a background. Part II will support the technique's rationale by illustrating how real IMS data for major government systems behaves over time. Part III will formally present the new technique, as well as provide examples and show results from actual contracts. Finally, Part IV will reflect on the technique proposed in this paper based on a third-party evaluation.

Part I: Background

The Integrated Master Schedule is more than a tool for managing the time phasing and resource allocation of project activities. Decision makers find value in an IMS for its ability to evaluate the effect of schedule risk and provide early warning for schedule slip. The information gives insight into one of the three categories for total contract performance risk, the other two being cost and technical risk. In general, realized risk in any one area can only be offset at the expense of performance in one of both of the others (see Figure 1 below).[1] For example, if the IMS predicts a significant schedule delay, the decision maker may allocate additional resources to the problem or reduce systems capabilities in order to maintain the contractual delivery date. Therefore, accurate and timely information on schedule risk provides flexibility in project management.

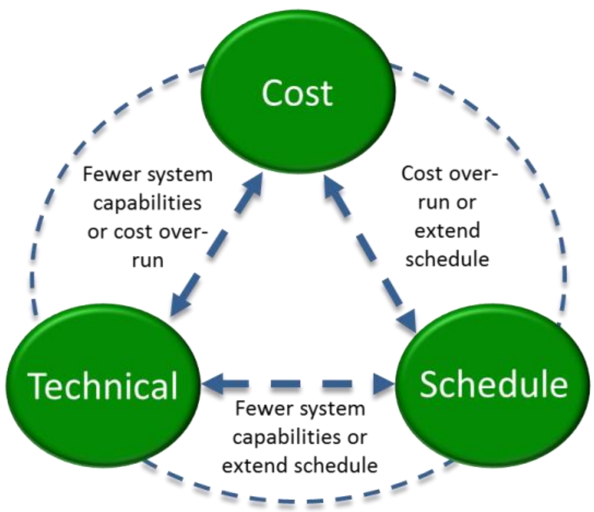


Figure 1: Cost, schedule and technical risk interrelationship

For many Major Defense Acquisition Program (MDAP) contracts, the IMS has a poor history of accurately reflecting schedule risk. In an analysis of 12 MDAP contracts,[2]

[1] There is often a positive correlation amongst cost, schedule and technical risk. Sometimes managers can only mitigate poor performance.

[2] MDAP contract data accessed through the Earned Value Management Central Repository (EVM-CR). This paper reflects an analysis of all contract schedules available. Data excluded: contracts without an approved Non-Disclosure Agreement (NDA); all contracts without data spanning contract start to end; IMSs in PDF or picture format; IMSs in Primavera (no active license). Currently 12 contracts with 133 schedule observations. Uniform data extraction methodology used across contract schedules: extracted 19 standard fields for non-summary activities; converted into Microsoft Excel flat-files and used standard

schedule slips were not apparent until late in the project. In Figure 2 below, one can see that the IMS largely does not indicate any delays until after half-way through the original schedule duration. In fact, as a project approaches its expected end date further delays are developed, creating a tail chase.[3] The situation gives decision makers minimal leeway for managing tradeoffs and implementing well-informed strategies.

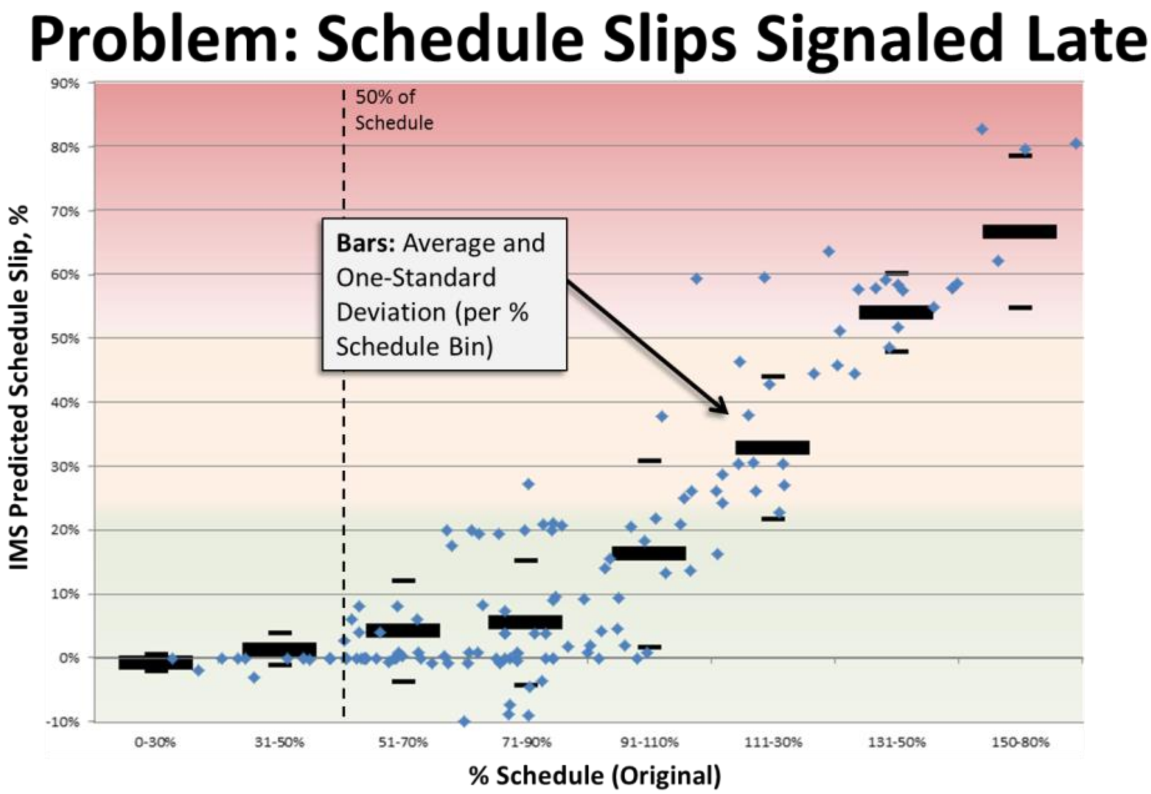


Figure 2: Predicted schedule slip reported by IMS submission

Because only the IMS itself can answer the question "when will the deliverable arrive?" a number of schedule metrics have been devised. For example, the Baseline Execution Index (BEI) indicates the efficiency with which tasks have been accomplished when measured against the baseline tasks.[4] However, the BEI is a backwards looking metric. It gives the number of total tasks finished compared to the number of tasks

---

template for metrics gathering and analysis. Additional contract insight gained from Defense Acquisition Management Information Retrieval (DAMIR).

[3] Cf. "Ending the EAC Tail Chase: An Unbiased EAC Predictor using Progress Metrics, Eric R. Druker, et al., SCEA/ISPA, 2007.

[4] DCMA, "Earned Value Management System (EVM) Program Analysis Pamphlet (PAP)," pp. 16-18.

planned, with no regard for relative task importance, lateness, or sequencing. The BEI does not provide a basis for risk-adjusted end dates; it can only indicate past task performance. Similar conclusions are reached for other metrics like task hit-miss, critical path length index, and ratios of high float or leads or hard constraints.

Guidance from authoritative sources, including the GAO Schedule Assessment Guide, the DCMA 14-Point Assessment, and the Joint Cost and Schedule Risk and Uncertainty Handbook, has traditionally focused on gauging the quality of the most recent IMS submission.[5] This is because the IMS is a "living document" where only the latest IMS incorporates up-to-date information. However, many still find it difficult to ascertain the realism of any given schedule because current schedule quality is a necessary but not sufficient criterion for schedule realism.

The inadequacy of status quo metrics stems from the perceived continuing irrelevancy of previous IMS submissions. The metrics are based on a trust that schedules have rigorously maintained a baseline, but none have been derived that will verify a rigorous baseline through cross-IMS analysis. Without a point of reference to ensure logical evolution, the current IMS can only say so much. It is important to understand, for example, what baseline changes have occurred over time and how actual performance has measured to near-term plans. While schedules are living documents, the initial baseline stands as the best available point of reference. This baseline is valid for three major reasons. First, planners tend to know the major activities involved in the execution of a project. All systems can be said to have historical analogies, even those considered revolutionary. Second, contractors generally have well-defined processes for developing these systems. Third, the IMS undergoes an Integrated Baseline Review (IBR) after which both the contractor and client agree to the plan and its efficacy. Thus, the IMS from its outset may be viewed by estimators primarily as a tool for measuring the scheduler's ability to plan in the near-term.

This paper argues that an activity's baseline from the initial IMS is relevant through subsequent submissions. Planners generally do a good job of laying out major activities, and so early performance on near-term activities should be a good indicator of total schedule realism. Practitioners have often heard anecdotally that early schedule slip

---

[5] A recent and important exception to this rule is GAO's "Best Practices for Project Schedules," released in Dec. 2015, which advocates a continuing analysis of the baseline schedule relative to the most recent instantiation of the schedule. It states in the 10th and final best practice: "The baseline schedule is not the same as the current schedule. The current schedule is updated from actual performance data." See GAO-16-89G. Rigorous maintenance of the schedule's baseline would create a fair amount of convergence between the predictions from the contractor's IMS and the predictions from the independent estimating technique that will be presented on the following pages. However, even if a schedule had a rigorously maintained baseline, it does not prevent the problem of optimism bias in forecasting task completion efficiency. The author's technique is not affected by the optimism (or pessimism) of forecasts in subsequent schedules; it is only affected by such factors as built into the baseline schedule.

cannot be made up despite managerial tactics. This paper will explore that notion and whether it finds support in the data. It will be shown that by tracing near-term activities through subsequent IMSs and comparing them to their original baseline, as opposed to the current baseline, one may extrapolate a more realistic contract finish date far earlier in the project.

Part II: IMS Observations and Relationships

This section will reveal facts and trends about IMS data across the lifecycle of a contract. It will be shown that, in general, projects rely on forecasts where expectations of task performance become increasingly detached from historical actuals. Not only are forecasts biased toward undue optimism, but schedule quality tends to decrease significantly over time leading to IMSs with relatively few useful interrelationships between activities. The charts that will be shown use a variety of x-axes. In normalizing schedule duration, there are three versions of % schedule used: original, current, and latest (see depiction below). The last (non-normalized) x-axis measure used is "# Months Past Initial IMS" in which data are binned according to the number of months the IMS was submitted from the very first IMS.

$$\%\ \text{Schedule}_x = \frac{(\text{Schedule 'As Of'} - \text{Contract Start}_0)}{(\text{Predicted End}_x - \text{Contract Start}_0)}$$

$$\text{where Predicted End}_x = \begin{cases} \text{First IMS (Original)} \\ \text{'As Of' IMS (Current)} \\ \text{Last IMS (Latest)} \end{cases}$$

The mean actual BEI values, binned by % schedule (current) in blue, is depicted in Figure 3 below. The corresponding mean forecast BEI is in red. BEI values of 1.0 represent completing the same number of tasks as had been planned by the submission's 'as of' date. Values above 1.0 represent working ahead of plan and value below 1.0 represent falling behind in task completion. That the average forecast task performance jumps well above the actuals strikes one immediately. Anticipated efficiencies increase dramatically even though actual performance tends not to stray far from 1.0. The spread between these two metrics mirror their cost counterparts, the Cost Performance Index (CPI) and the To-Complete Cost Performance Index (TCPI). When forecasted values lay well above their historical actuals, there is significant risk as the project schedule relies on unreasonable task performance.

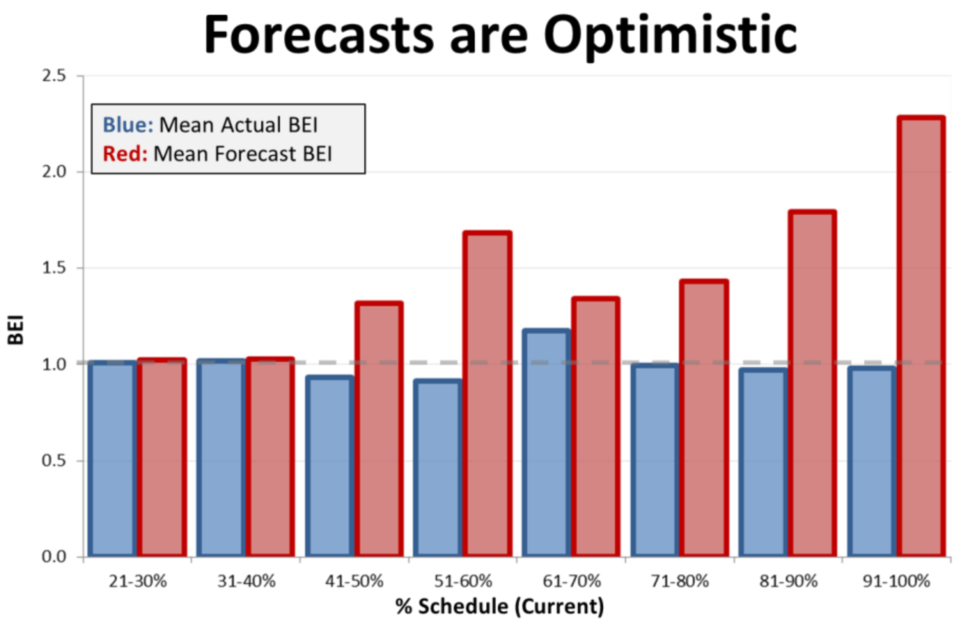


Figure 3: Actual BEI (blue) and Forecasted BEI (red). BEI = 1.0 denotes working to plan.

Though schedules rely on completing tasks on time, the BEI does not give a holistic measure of efficiency. As stated earlier, it does not measure task lateness or importance, but rather aggregate completions irrespective of sequencing. Additionally, as the project progresses, the weight of past activities starts to drive the BEI metric. For example, a project has completed 150 tasks compared to 200 planned, the BEI is 0.75. However, if 450 tasks were completed to 500 planned, the BEI is 0.90. The BEI, then, trends towards 1.0 even though the contractor remains behind at a constant rate of 50 tasks.[6]

[6] A similar situation is true of cost metric counterparts like cumulative Cost Performance Index (CPI) and Schedule Performance Index (SPI).

Task performance may also be readily evaluated by counting the number of tasks which finished late relative to their baseline dates. Figure 4 below shows just that: the percentage of discrete tasks which finished late to baseline over the 9 months prior to submission 'as of' date. This sample shows task performance degradation over time. The interpretation of a linear slope is that for every 2% of schedule that passes, an additional 1% of recently finished tasks will finish late to baseline. This observation strengthens the case that future schedule optimism is unfounded. Rather than completing tasks more efficiently in the future, reality suggests that more tasks than not will finish late.

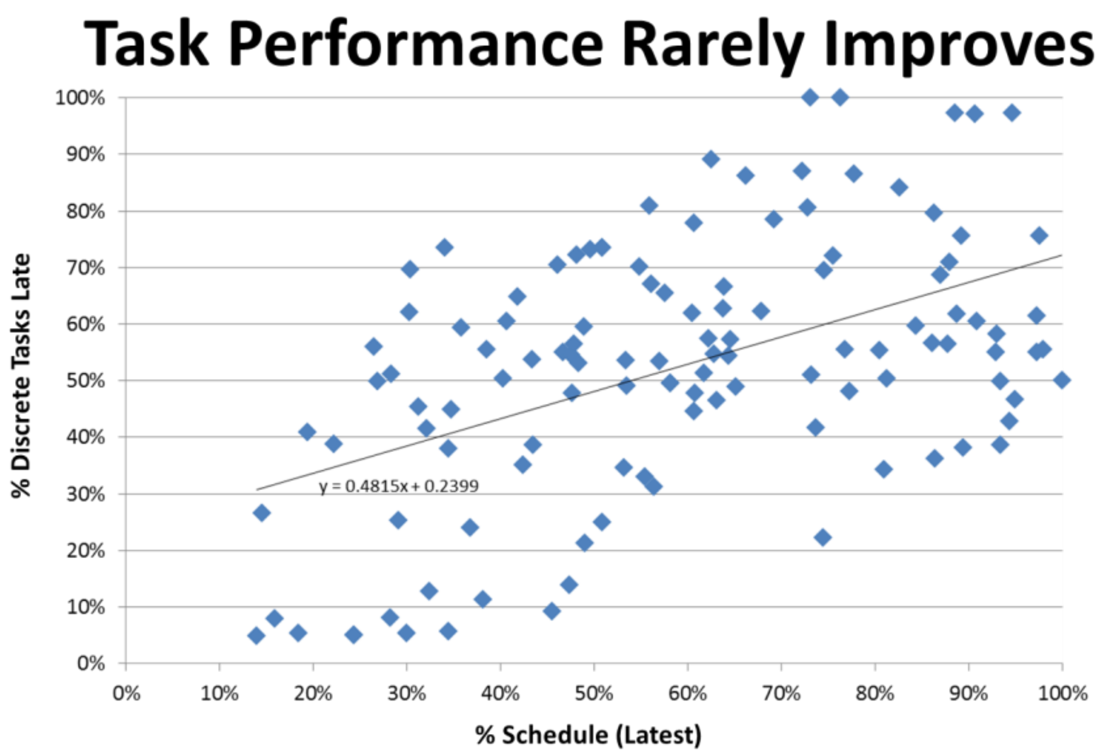


Figure 4: % Discrete Tasks which finish late to current baseline over past 9 months

Thus far we have been seeing tasks performance relative to their current baseline. But how has that baseline changed over time when compared to the very first IMS? For each change, has the client approved the baseline change? The blue bars in Figure 5 below show the average percentage of activities with a change in baseline finish date. Of activities which persist through IMS submissions, roughly a quarter of them have changes to their baseline. On average, those activities with baseline changes eventually slip over 2 months (red scatter, right axis) and in several cases well over a year (not depicted).

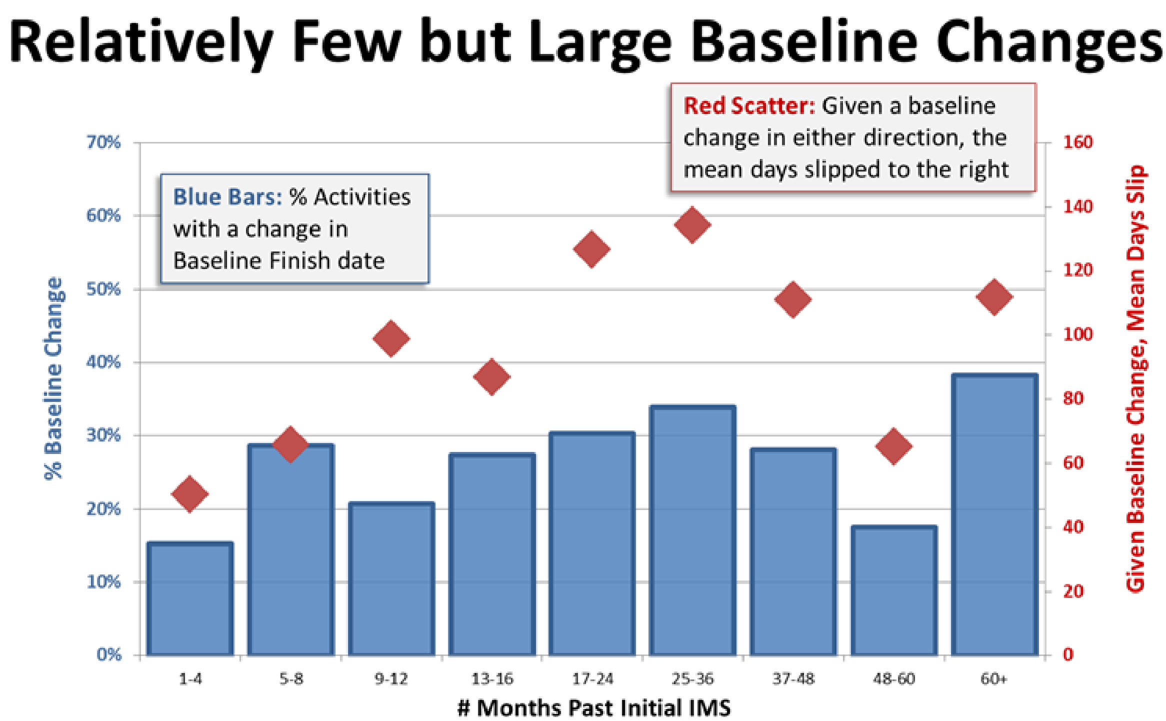


Figure 5: % Baselines change (blue, left axis). Mean days change of new baseline (red, right axis).

Aside from task performance, another key metric used to measure schedule quality is the occurrence of constraints. The DCMA 14-Point Assessment flags any IMS with more than 5% hard constraints as they prevent tasks from being moved by their dependencies and, therefore, prevent the schedule from being logic-driven.[7] Figure 6 below shows that many schedules constrain more than 5%, and up to 20%, of activities (blue). Of those activities with constraints, generally 50% or more are binding. Binding means that the forecast start or end date equals the constraint date. When a constraint binds an activity, it is not driven by its predecessors and may no longer slip. Though an activity may have a sound reason for requiring a hard constraint, such tasks can distort the IMS and mask a probable schedule slip.

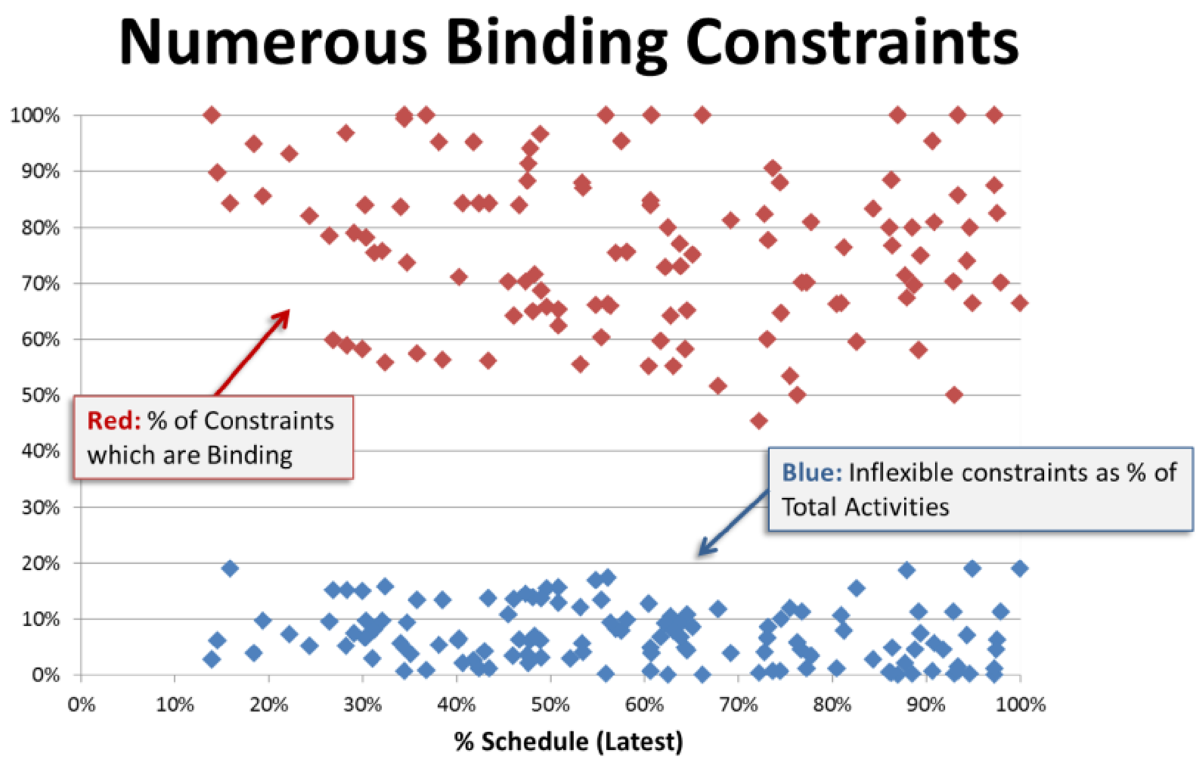


Figure 6: (Blue) % of total activities constrained. (Red) % of constrained activities which are binding.

[7] DCMA, "Earned Value Management System (EVM) Program Analysis Pamphlet (PAP)," pp. 16-27.

Maybe the most important concept in schedule construction and maintenance is logic: the idea that every activity must have at least one predecessor and one successor activity. Logic drives the sequencing of activities and interconnects the schedule so that a change to any one given activity has a ripple on all successors. The DCMA 14-Point Assessment also deems it risky for more than 5% of incomplete activities to have missing logic in an IMS.[8] Figure 7 below depicts the number of tasks missing logic forecasted to finish over the next three months. It is apparent that more than 5% from this subset of tasks do not have logic links. Additionally, the percentage of tasks without logic tends to grow over time with certain schedules missing up to 80%. This indicates that schedule quality tends to decrease over time as an increasing number of near-term tasks are not tied to the schedule logic.

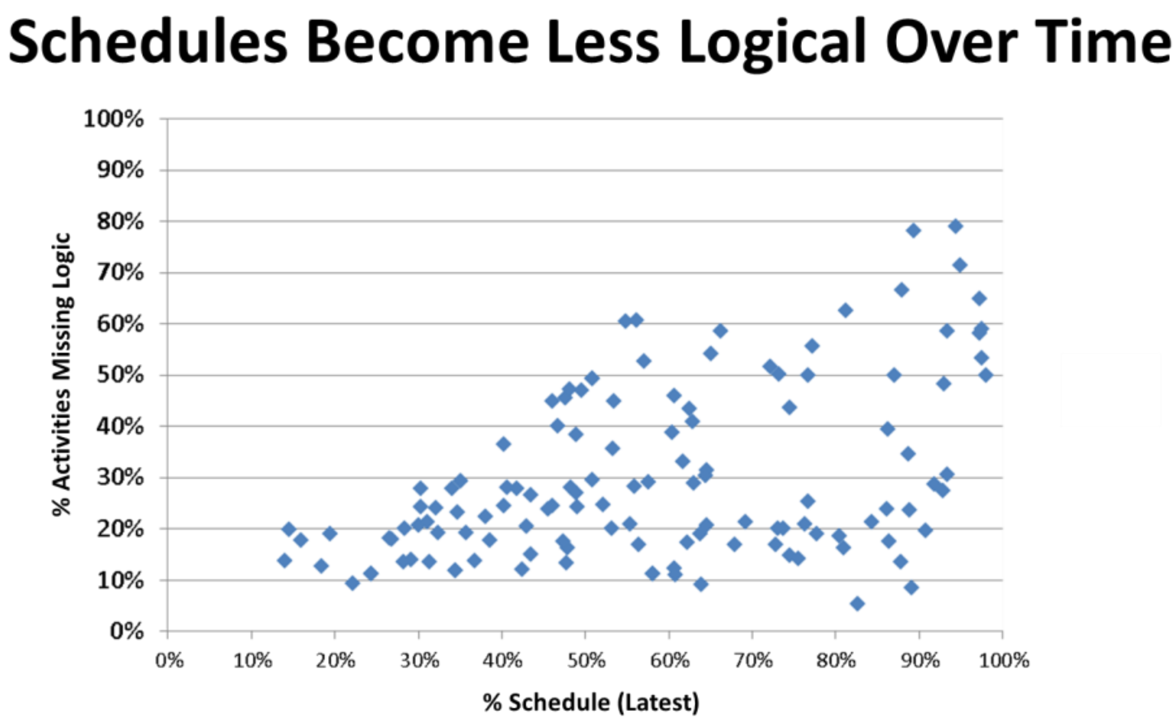


Figure 7: % of activities missing both predecessor & successor logic links in next 3 months

The preceding section has shown that, in general, schedule quality and performance tend to decrease over the course of a contract. A high number of binding constraints and increasing number of late tasks leads to schedule optimism. Unapproved baseline

[8] DCMA, "Earned Value Management System (EVM) Program Analysis Pamphlet (PAP)," pp. 25.

changes can conceal poor performance. Fewer logic links indicate a poorly maintained schedule. These generalizations have a negative impact on schedule realism. Even a high quality rating for any given IMS does not mean that it has also evolved in a consistent manner. The relevant question becomes "can current IMS data be better utilized to measure schedule risk and extrapolate a realistic end date far earlier in the project?" Part III of this paper will explore a process for doing just that.

Part III: An Improved Schedule Estimator

The lynchpin of the proposed schedule estimating technique is the baseline IMS. The first available IMS, preferably the one immediately following IBR, will be used to independently track activities over subsequent submissions. Activity end dates are traced over time using their designated activity names and/or unique IDs. The updated end dates will then be compared to their original baseline end date and total float (slack days).[9] This process disregards: new activities which get built into the schedule over time; activities which change identifiers; and changes to activity leads, lags, constraints, and sequencing. While the omissions appear concerning, remember, the primary purpose is to trace performance of the relatively near-term baseline plan. Both contractual parties agreed to the baseline IMS which took months to develop. That some activities cannot be traced (e.g. a task broken up into three new, smaller tasks) is simply a reality insofar as there is no record of the changes.

In order to determine schedule slip, the analyst looks across all tasks still identifiable in subsequent IMS submissions and compares them to their original baseline. The activity which has slipped most to that baseline, with the original days of float factored in, drives the independently predicted end date. That number of days slip is added to the original contract closeout date.

$$\text{Predicted End Date}(t_n) = \text{Baseline End Date}(t_0) + \text{Slip}(t_n)$$
$$\text{Slip}(t_n) = \max(\{f(x_i): i = 1,\ldots,k\})$$
$$\text{where } f(x_i) = [\text{Current Finish}_i - \text{Baseline Finish}_i - \text{Baseline Total Float}_i]$$
$$\text{and } \max(\{f(x_i): i = 1,\ldots,k\}) \geq 0$$
$$x_i = \text{observed activity}$$

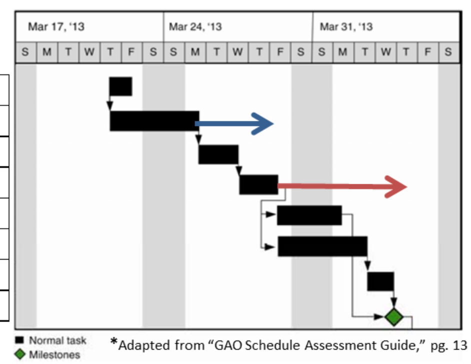
9 Total float (also called total slack) is the number of days an activity is allowed to move to the right before any further slips affect the project end date. When an activity has zero days of slack, it might find itself on the "critical path," or a sequence of activities representing the shortest duration between the schedule's 'as of' date and the end of the project.

When the schedule has evolved to the extent that the current activities no longer reflect the baseline, the predicted end date stabilizes and the estimate is determined. This occurs because the tasks in the baseline IMS have their finish dates realized and the planning packages open up into new work packages. The rule implemented by the author is to stop the analysis when more than 95% of the activities from the original baseline are either finished or no longer appear in the current schedule. The final prediction for contract end is then set. An example will illustrate this methodology.

The first IMS sets the baseline for future activities. All activity finish dates will be evaluated for each subsequent IMS and compared to their baseline. In the case of IMS #2, several activities have experienced slips in their forecast finish dates. However, "Frame first floor walls" has slipped 4 days while it only had 3 days of float available. This means that the task will affect the start of all of its successor tasks and push out the milestone "Framing complete" by 1 day, *ceteris paribus*. In IMS #3, "Install roof decking" has slipped 7 days to baseline, and factoring in its original 3 days of float, should now affect "Framing complete" by 4 days.

An Example

	IMS #1		IMS #2		IMS #3	
Activity Name	Baseline Finish	Total Float	Finish	Slip	Finish	Slip
Rough grade property	03-20-13	3	03-20-13	0	03-20-13	0
Frame first floor walls	03-25-13	3	03-29-13	1	03-29-13	1
Install roof trusses	03-27-13	3	03-29-13	0	03-29-13	0
Install roof decking	03-29-13	3	03-31-13	0	04-05-13	4
Rough-in framing inspection	04-01-13	5	04-03-13	0	04-03-13	0
Form and pour driveway	04-02-13	3	04-04-13	0	04-05-13	0
Finish excavate and pour garage	04-03-13	3	04-05-13	0	04-06-13	0
Framing complete	04-03-13	3	04-05-13	0	04-06-13	0
			Max Slip:	1	Max Slip:	4



*Adapted from "GAO Schedule Assessment Guide," pg. 13

Figure 8: A simplified example of an evolving IMS

How has "Framing complete" been able to stay on track? In the example above, the forecasted tasks have had their durations implicitly squeezed in order to make up time. In reality, forecasted discrete tasks do not regularly show shorter durations than past tasks. A plausible reason is that the long-term future is represented by planning packages which incorporate many tasks. When opened, the resulting tasks are given realistic durations and distribute the implicit duration squeeze onto planning packages still further out. Although evidence cannot be concretely pulled from the data, it has been shown that forecasted performance is optimistic and there are numerous binding constraints. Such behavior is sufficient to maintain a fixed near-term schedule, even if it is unrealistic in the long-term. The total schedule might be saved, however, because of the future's vagueness. As time progresses, the pool of undetailed tasks dwindles and increasingly focuses implicit duration compression. By the time the schedule's unrealism is noted (possibly by pure reckoning), it would be relatively late in the project.

The proposed technique disregards over-optimism by relying on early task performance to baseline. Because major activities are reasonably phased from the start, the total schedule receives a one-for-one slip with the near-term (realized) schedule. The idea becomes increasingly attractive when considering the fact that schedules tend to lose an alarming number of logic links over time. This means tracking near-term performance late in the schedule tells one little, since relatively few tasks are imparted with interrelationships. The estimation technique is not intended to project exact dates, but should be viewed as a rough order of magnitude for would-be schedule slips. The results of this technique are presented for 8 MDAP contracts in a standard chart template.[10]

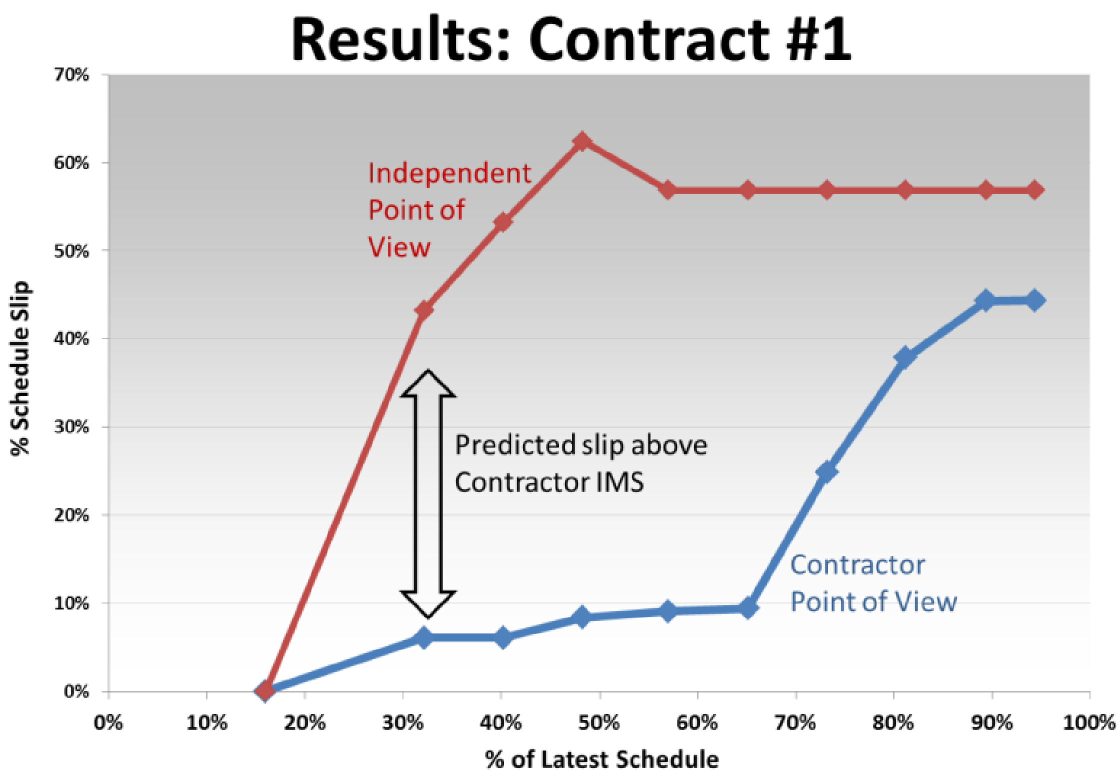


Figure 9: Results from an actual contract

Above, one may see that from the contractor's point of view, the IMS shows less than 10% predicted schedule slip up until near 70% of the latest schedule. Then significant delays are realized, and the project eventually slips by almost 45% to baseline. The

[10] Though 12 MDAP contracts were collected and analyzed, 4 were of the type Indefinite Delivery, Indefinite Quantity (IDIQ). The 4 IDIQ contracts are not shown because new contract task orders or scope changes have a clear effect on schedule duration. More data insight is needed to compensate for the volatility in scope.

independent point of view, representing the technique proposed in this paper, quickly registers schedule risk to within about 10% of actual slip. Note that the (red) independent line plateaus shortly after 50% of latest schedule. This is where the current schedule largely reflects new activities; in other words, activities from the baseline IMS have for the most part been completed.

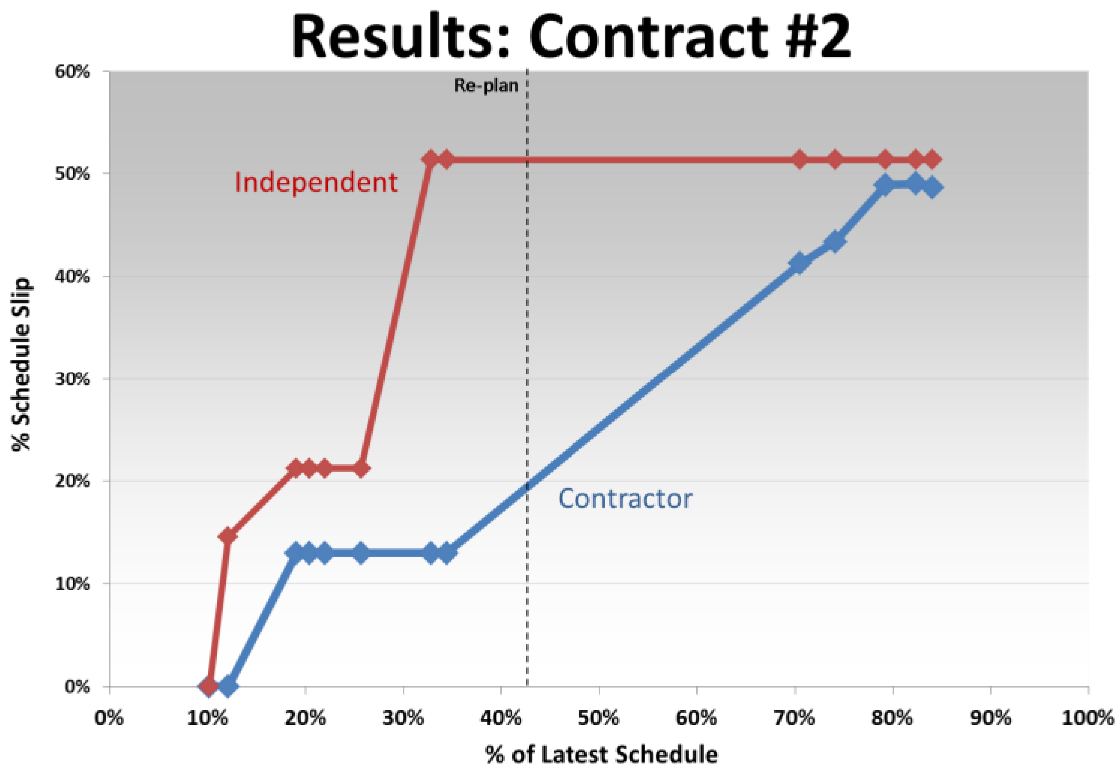


Figure 10: Results from another actual contract

Results from a second contract are similar to the first: the independent technique gives a good approximation of the realized schedule slip fairly early on in the contract. This contract, however, has a major re-plan shortly after 40% of latest schedule.[11] Because the independent line plateaus first, it has the expected duration of the slip associated with the re-plan embedded in it. This is valid because the project did not increase quantity or add significant capabilities. Additional budget and schedule were allocated to a struggling project. This method is useful where re-plans are an admission of *de facto* poor performance to plan because of technical problems or otherwise. This method does not do a good job in its raw form presented here for Indefinite Delivery, Indefinite Quantity (IDIQ) contract types.

[11] Re-plans are ascertained by inspecting the Contract Performance Reports (CPRs) Formats 1-5 as well as contract narratives from Defense Acquisition Management Information Retrieval (DAMIR).

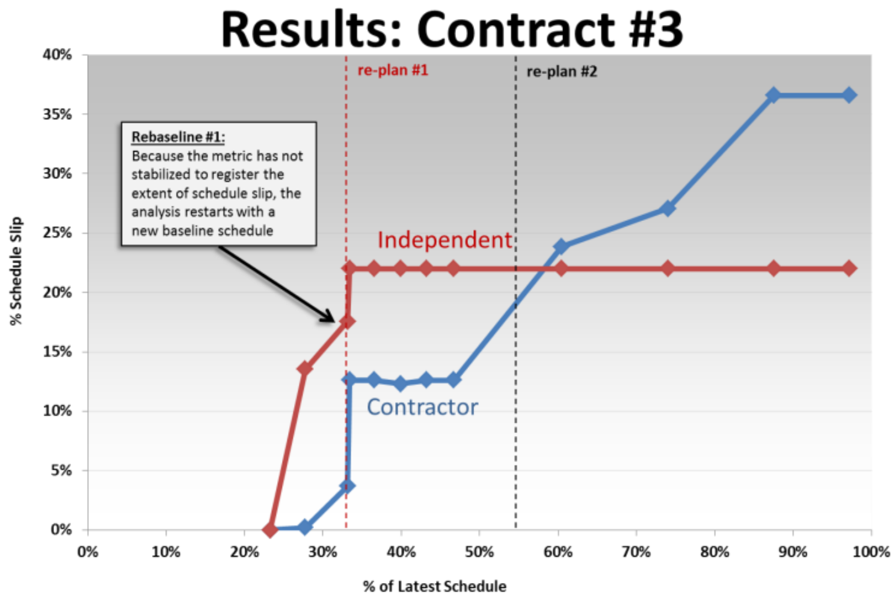


Figure 11: Results from a third actual contract

The contract in Figure 11 above shows a situation where there was a re-scope early in the contract. Because a new plan was implemented before the independent metric was able to measure the realized slip in near-term baseline tasks, it was unable to register the full extent of the total schedule slip. In such situations, the re-planned IMS may be used as a fresh baseline, and the analysis restarts. The contract above underwent a second major re-plan shortly thereafter.

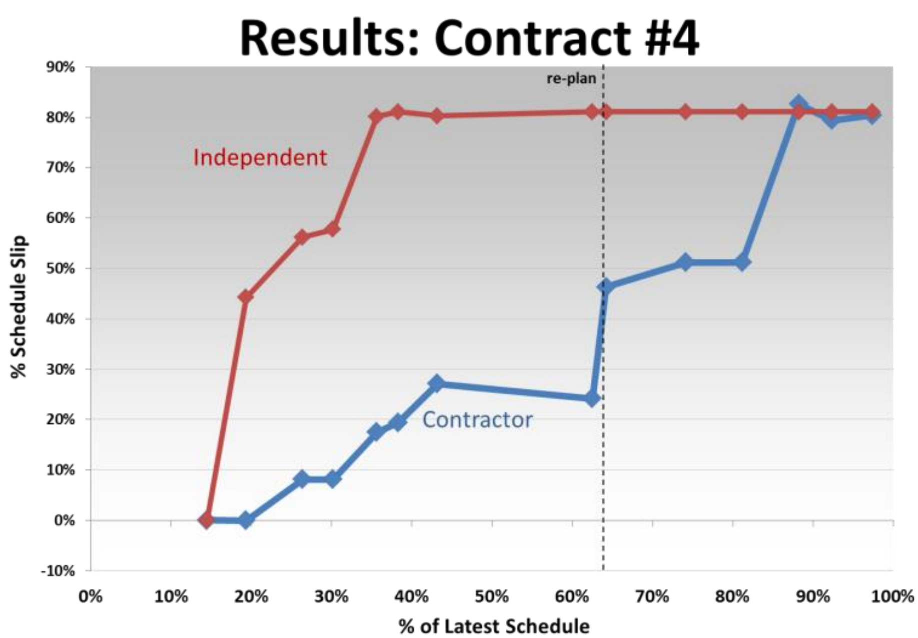


Figure 12: Results from a fourth actual contract

In Figure 12 above, it is again observed that early in the contract, the independent metric closely forecasts the true schedule slip. In this instance, the technique quickly shows 40% then 80% slip. Though the project eventually conformed to that prediction, it brings up the problem of the self-fulfilling prophecy. If a decision-maker admits to significant delays early on, then the project will be managed to a longer timeframe and the personnel won't feel schedule pressure. This could, in fact, lead to further delays down the line. Results for the remaining 4 contracts are shown in Figure 13 below.

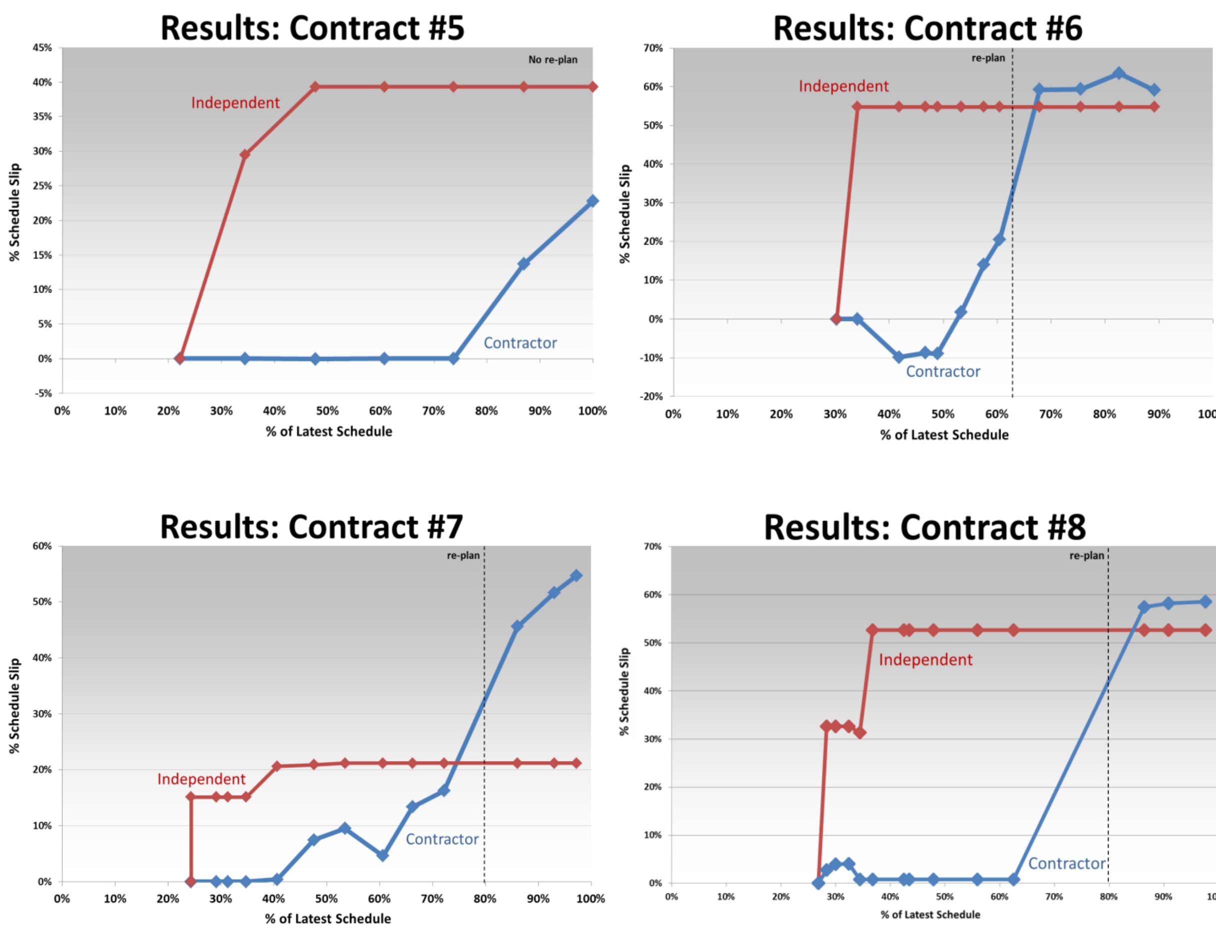


Figure 13: Results from actual contracts five through eight

The independent metric finds a useful application in cost estimation as well. If the metric predicts a one year slip over that reported in the IMS, the analyst may extend the average project cost burn rate out for the same duration. Figure 14 below does just that using one of the real contracts analyzed in this paper. The red line is the mean actual burn rate multiplied by the number of months slip the independent metric predicts over the IMS plus the cost of the traditional Independent Estimate At Completion (IEAC). The new independent cost estimate acts similar to the schedule estimator and gives a good early indication of realized costs.

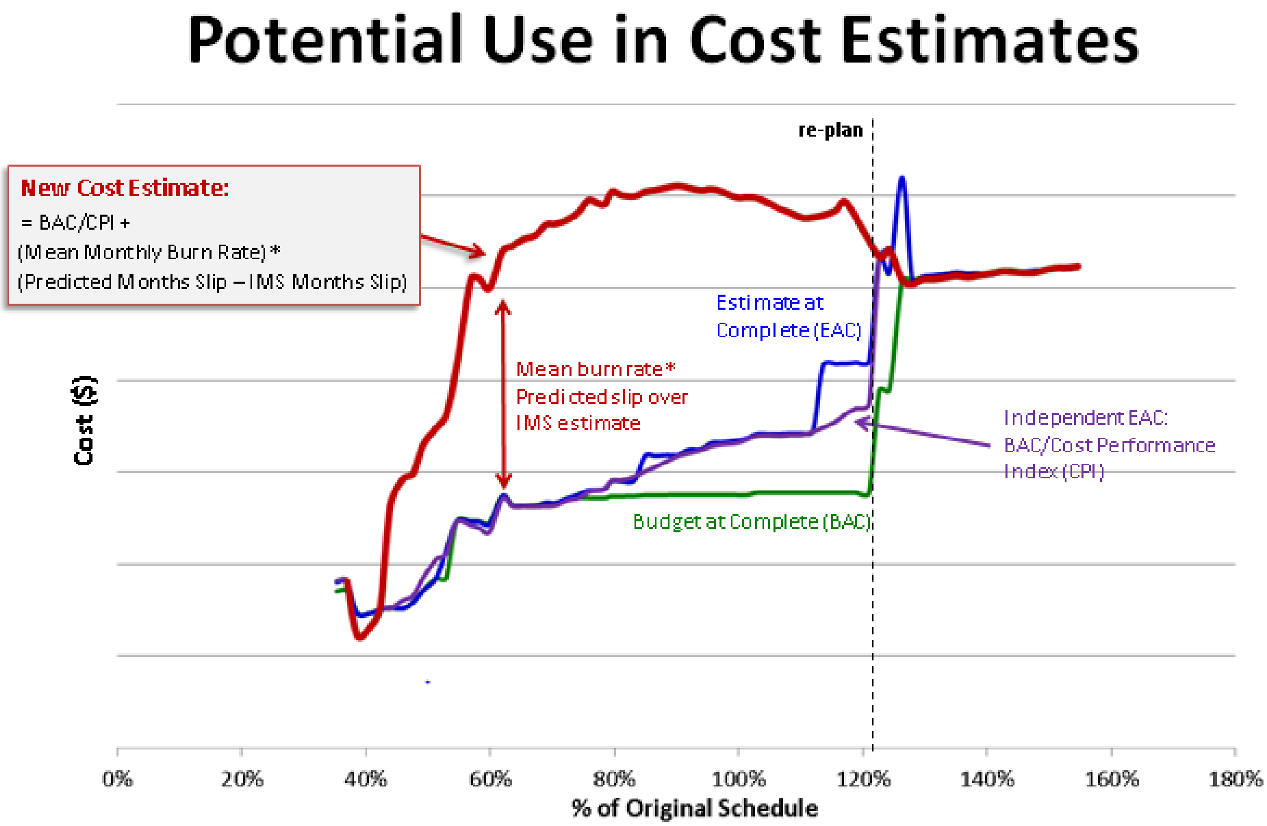


Figure 14: Application of independent metric to cost estimation for an actual contract

How may an estimator use this metric to improve outcomes on MDAP contracts? First and foremost, it gives the decision maker an early indicator to the magnitude of a potential schedule slip. Figure 15 below illustrates the new metric's predictive power. By 50% of the latest schedule duration, the independent metric registers a majority of realized schedule slip. The mean absolute error for the 41-50% of schedule bin is only seven months; compare that the contractor's 25 months of error. It equates to a year-and-a-half of schedule slip on average not yet realized by the contractor's IMS. In fact, the contractor IMS did not reach a mean absolute error of seven months until close to 80% of the total schedule had elapsed. The reduced forecasting error provided far earlier in the project allows management to plan early for cost-schedule-technical tradeoffs. Early recognition of catastrophic difficulties also leads to greater flexibility in project termination because of fewer costs sunk. As shown above, the technique may also prove useful in providing more realistic cost estimates. Finally, the method allows for traceability between submissions to verify schedule quality. For example it is useful to understand how volatile activities are and whether the IMS is continually being re-invented. Schedule quality as traditionally calculated largely ignores the quality of the schedule's evolution.

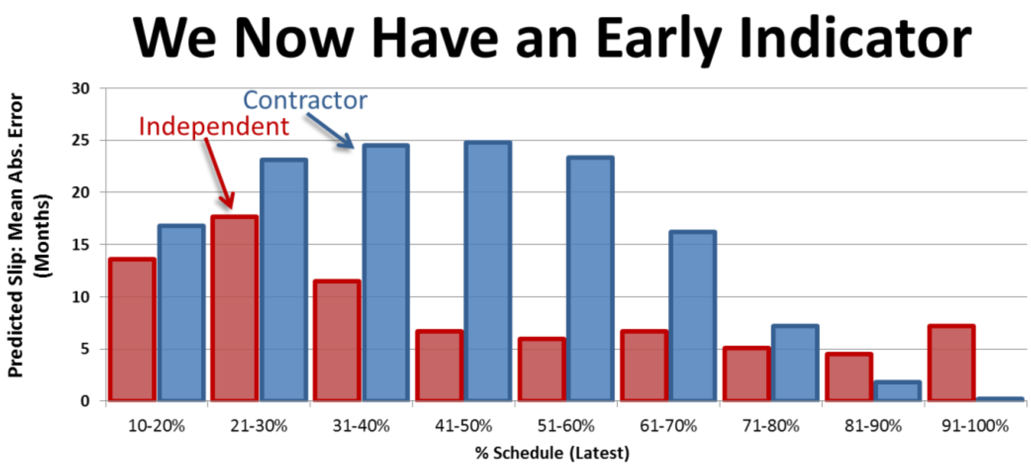


Figure 15: Mean absolute error of contractor and independent schedule estimates

Using the new technique presented here, an analyst may catch potential schedule slip earlier in the project with relative accuracy. While schedules develop over time, planners do a good job at the outset of phasing major milestones and near-term work packages. Though projects change, the deliverable doesn't often transform completely. However, cumulative changes to a schedule tend to deteriorate its realism. As has been shown, schedule performance is best registered early for a variety of reasons including increasingly optimistic forecasts and decreasing quality. It may be concluded that contractors quickly reveal their pace of work and "settle" into a performance; or management strategies such as work-arounds can do little to alleviate ailing projects. Until a culture of improved schedule maintenance takes root, performance to baseline early in the project may serve as the best indication of realized schedule slip.

Part IV: Reflections

A year after the development of the estimating technique presented in this paper, Shedrick Bridgeforth independently tested it along with various other schedule estimating techniques.[12] He found that for a set of 12 completed defense satellite development efforts, the technique presented in this paper, which he called the Independent Duration Estimate (IDE), outperformed all other EVMS-based approaches in predicting the realized schedule. Bridgeforth wrote that "These results suggest Lofgren's approach (IDE) is the most accurate technique." In particular, the so-called IDE approach provided more realistic estimates far earlier in the development cycle than other techniques. Bridgeforth recommends using the IDE technique for space development contracts when the data are available.

Figure 16 below shows Bridgeforth's results for a variation of the IDE model (red) compared to an estimate based on the Schedule Performance Index (blue).[13] Notice how the IDE model begins registering schedule slip, and minimizing the error, well in advance of the status quo metric. The advanced warning amounts to approximately 20-30% of realized schedule, often representing more than a year of lead time for managers of satellite development contracts. The behavior is consistent with the results in this paper shown in Figure 15 above.

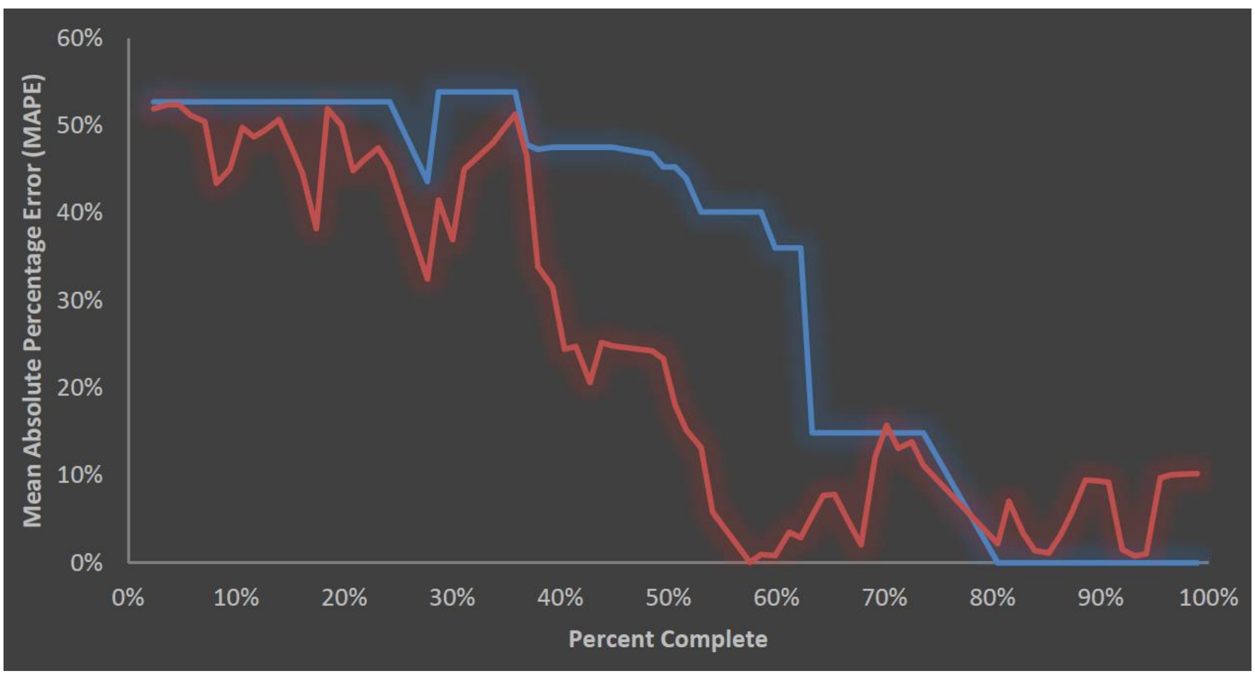


Figure 166: Results from Bridgeforth (2015) showing performance of the IDE technique from this paper (red) and the status quo (blue)

---

[12] Bridgeforth, Shedrick. 11 Sep. 2015. "Using Earned Value Data to Forecast the Duration of Department of Defense Space Acquisition Programs." *Journal of Cost Analysis and Parametrics*, Vol. 8, Issue 2, pp. 92-107.

[13] Bridgeforth, Shedrick. "Using Earned Value Data to Forecast the Duration of Department of Defense Space Acquisition Programs." Thesis for Air Force Institute of Technology (AFIT). http://www.dtic.mil/docs/citations/ADA615411, pp 71.

While Bridgeforth's results support the conclusions in this paper, he cautions that "Lofgren's IDE framework does not consider the critical path; it considers all tasks as equally important. The IDE may struggle to become a best practice because it ignores the CPM [Critical Path Method] and is relatively new." While the technique is both new and not well recognized in the cost and schedule analysis communities, it is not true that the technique ignores the CPM. As explained on in Part III, the total float (slack) from the baseline schedule is taken into account, factoring in the logic behind the baseline critical path. Activities in successive IMSs that slip more days than they had total float in the baseline schedule then compete for an unofficial critical path to become schedule drivers. The technique presented here is independent precisely because it ignores new activities and changes to leads, lags, constraints, and sequencing that may alter the critical path. Similarly, it ignores increasingly optimistic forecasts of task performance which may artificially compress the critical path.

If the activities and relationships networked in the baseline schedule remained constant over time, the activities driving the IDE technique should be exactly those activities found on the contractor's critical path in the IMS. What distorts the true signal of schedule progress is the cumulative changes to the content and interrelationships of activities. By stripping away such "noise," which in many cases results from efforts to mitigate realized risks, the technique proposed here focuses in on actual performance to baseline plan for near-term tasks. It rejects updated expectations to forecasts of task efficiency, which often has an optimistic bias.

As the single source of comprehensive schedule information on most projects, the IMS's predictive ability rests on its quality. It is preferred that the IMS exhibits high quality such that the analyst can develop actionable plans using its forecast. Currently, analysts spend a great deal of effort trying to tease out the biases in schedules to generate more "realistic" forecasts. The technique presented in this paper is no different. It searches for bias by independently tracking baseline activities through subsequent schedules. All such manipulation of native schedule data to support independent predictions is highly speculative. Attention to schedule maintenance is a superior use of scarce resources relative to guessing its inherent biases. Longitudinal checks of the sort described in this paper, particularly Part II, would actually expose the need to eliminate the biases that analysts attempt to exploit for the purpose of prediction. Oversight of this sort can go a long way to ensuring high schedule quality from the outset.

Though improving schedule quality is the first-best solution, in today's scheduling environment there is much to be done using the author's schedule estimating technique, including further refinements. First and foremost is the need for an expanded data set to repeatedly test the metric's robustness. Second, it important to scrutinize the tasks that drive the independent technique's estimate. For example, there may be an outlier that is not representative of actual schedule performance. In such cases, the technique may be better served by taking the median slip of the top five, ten, or twenty tasks. Third, to assess the value of this technique for IDIQ contracts, a more detailed approach is required, such as analysis by task-order. Further assessment is also needed for augmentation to cost estimation. One method (extending mean burn rates) has been presented in this paper, but many other approaches are possible. Finally, a large collection of IMSs used for analysis can also provide data-driven generalizations of schedule quality and realism over time. Additional insights into IMS evolution can help practitioners understand how to better manage their schedule.

Bibliography:

"Earned Value Management System (EVMS) Program Analysis Pamphlet (PAP)." Defense Contract Management Agency (DCMA). July 2012. http://www.dcma.mil/policy/200-1/PAM-200-1.pdf. Accessed 10 April 2014.

"Ending the EAC Tail Chase: An Unbiased EAC Predictor using Progress Metrics, Eric R. Druker, et al., SCEA/ISPA, 2007.

"GAO Schedule Assessment Guide." U.S. Government Accountability Office (GAO). 30 May 2012. http://www.gao.gov/assets/600/591240.pdf. Accessed 10 April 2014.

"IMS DID." Office of the Undersecretary for Defense Acquisition, Technology, & Logistics Performance Assessments and Root Cause Analyses (OUSD AT&L PARCA). 30 March 2005. http://dcarc.cape.osd.mil/EVM/Documents.aspx. Accessed 10 April 2014.

"IPMR Implementation Guide." Office of the Undersecretary for Defense Acquisition, Technology, & Logistics Performance Assessments and Root Cause Analyses (OUSD AT&L PARCA). July 21 2012. http://dcarc.cape.osd.mil/EVM/Documents.aspx. Accessed 10 April 2014.

"Joint Cost and Schedule Risk and Uncertainty Handbook (CSRUH)." Naval Center for Cost Analysis (NCCA). 16 July 2013. https://www.ncca.navy.mil/tools/csruh/index.cfm%20. Accessed 10 April 2014.